



# Low-rank matrix recovery with Ky Fan 2- $k$ -norm

Xuan Vinh Doan<sup>1,2</sup> · Stephen Vavasis<sup>3</sup>

Received: 14 March 2020 / Accepted: 1 May 2021 / Published online: 2 June 2021  
© The Author(s) 2021

## Abstract

Low-rank matrix recovery problem is difficult due to its non-convex properties and it is usually solved using convex relaxation approaches. In this paper, we formulate the non-convex low-rank matrix recovery problem exactly using novel Ky Fan 2- $k$ -norm-based models. A general difference of convex functions algorithm (DCA) is developed to solve these models. A proximal point algorithm (PPA) framework is proposed to solve sub-problems within the DCA, which allows us to handle large instances. Numerical results show that the proposed models achieve high recoverability rates as compared to the truncated nuclear norm method and the alternating bilinear optimization approach. The results also demonstrate that the proposed DCA with the PPA framework is efficient in handling larger instances.

**Keywords** Rank minimization · Ky Fan 2- $k$ -norm · Matrix recovery

## 1 Introduction

Matrix recovery problem concerns the construction of a matrix from incomplete information of its entries. This problem has a wide range of applications such as recommendation systems with incomplete information of users' ratings or sensor localization problem with partially observed distance matrices (see, e.g., [4]). In these applications, the matrix is usually known to be (approximately) low-rank. Finding these low-rank matrices are theoretically difficult due to their non-convex properties. Computationally, it is important to study the tractability of these problems given the large scale of datasets considered in practical applications. Recht et al. [21] studied the low-rank matrix recovery problem using a convex relaxation approach which is tractable. More precisely, in order to recover a low-rank matrix  $X \in \mathbb{R}^{m \times n}$  which

---

This work is partially supported by the Alan Turing Fellowship of the first author.

---

✉ Xuan Vinh Doan  
Xuan.Doan@wbs.ac.uk  
Stephen Vavasis  
vavasis@uwaterloo.ca

<sup>1</sup> Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom

<sup>2</sup> Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB, United Kingdom

<sup>3</sup> Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada

satisfies  $\mathcal{A}(X) = \mathbf{b}$ , where the linear map  $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$  and  $\mathbf{b} \in \mathbb{R}^p$ ,  $\mathbf{b} \neq \mathbf{0}$ , are given, the following convex optimization problem is proposed:

$$\begin{aligned} \min_X & \|X\|_* \\ \text{s.t. } & \mathcal{A}(X) = \mathbf{b}, \end{aligned} \quad (1)$$

where  $\|X\|_* = \sum_i \sigma_i(X)$  is the nuclear norm, the sum of all singular values of  $X$ . Recht et al. [21] showed the recoverability of this convex approach using some restricted isometry conditions of the linear map  $\mathcal{A}$ . In general, these restricted isometry conditions are not satisfied and the proposed convex relaxation can fail to recover the matrix  $X$ .

Low-rank matrices appear to be appropriate representations of data in other applications such as biclustering of gene expression data. Doan and Vavasis [7] proposed a convex approach to recover low-rank clusters using dual Ky Fan 2- $k$ -norm instead of the nuclear norm. Ky Fan 2- $k$ -norm is defined as

$$\|A\|_{k,2} = \left( \sum_{i=1}^k \sigma_i^2(A) \right)^{\frac{1}{2}}, \quad (2)$$

where  $\sigma_1 \geq \dots \geq \sigma_k \geq 0$  are the first  $k$  largest singular values of  $A$ ,  $k \leq k_0 = \text{rank}(A)$ . The dual norm of the Ky Fan 2- $k$ -norm is denoted by  $\|\cdot\|_{k,2}^*$ ,

$$\begin{aligned} \|A\|_{k,2}^* &= \max_X \langle A, X \rangle \\ \text{s.t. } & \|X\|_{k,2} \leq 1. \end{aligned} \quad (3)$$

These unitarily invariant norms (see, e.g., Bhatia [3]) and their gauge functions have been used in sparse prediction problems [1], low-rank regression analysis [8] and multi-task learning regularization [12]. When  $k = 1$ , the Ky Fan 2- $k$ -norm is the spectral norm,  $\|A\| = \sigma_1(A)$ , the largest singular value of  $A$ , whose dual norm is the nuclear norm. Similar to the nuclear norm, the dual Ky Fan 2- $k$ -norm with  $k > 1$  can be used to compute the  $k$ -approximation of a matrix  $A$  (Proposition 2.9, [7]), which demonstrates its low-rank property. Motivated by this low-rank property of the (dual) Ky Fan 2- $k$ -norm, which is more general than that of the nuclear norm, and its usage in other applications, we aim to utilize this norm to solve the matrix recovery problem.

In addition to the convex relaxation approach using nuclear norm, there have been many *non-convex* approaches for the matrix recovery problem. The recent survey paper by Nguyen et al. [18] surveys and extensively tests many of these methods. According to [18], a natural dichotomy distinguishes methods for which  $k$  is *a priori* unknown versus methods that prescribe  $k$ . Our proposed approach falls into the latter category. Among those methods that prescribe  $k$ , the next dichotomy is between methods based on nuclear-norm minimization versus those that attempt to directly minimize the residual  $\|\mathcal{A}(X) - \mathbf{b}\|$  by carrying out some form of constraint-preserving descent on  $X$ . Methods based on nuclear-norm minimization generally have better recovery rates according to the experiments of [18].

Although our method is in neither of these categories, it is more similar to nuclear-norm based methods of [18] because it minimizes convex norms and does not maintain feasible iterates. Of the non-convex methods, the method with the best recovery rate according to [18] is called *truncated nuclear norm minimization (TNN)* by Hu et al. [10]. TNN minimizes the non-convex function  $\|X\|_* - \|X\|_{k,1}$ , where  $\|X\|_{k,1}$  denotes the sum of the largest  $k$  singular values of  $X$ , the Ky Fan 1- $k$ -norm. It should be observed that  $\|X\|_* - \|X\|_{k,1}$  is nonnegative and is the sum of singular values  $k + 1, k + 2, \dots, \min(m, n)$ . Therefore,

$\|X\|_* - \|X\|_{k,1} = 0$  if and only if  $\text{rank}(X) \leq k$ , hence TNN is an exact formulation of the original problem. We are going to compare our proposed approach with TNN among others.

## 1.1 Contributions and paper outline

In this paper, we focus on non-convex approaches for the matrix recovery problem using the (dual) Ky Fan 2- $k$ -norm. Specifically, our contributions and the structure of the paper are as follows:

1. We propose a Ky Fan 2- $k$ -norm-based non-convex approach to solve the matrix recovery problem and discuss the proposed models in detail in Sect. 2.
2. We develop numerical algorithms to solve those models in Sect. 3, which includes the framework for difference of convex functions algorithm (DCA) and a proximal point algorithm (PPA) framework to solve the (dual) Ky Fan 2- $k$ -norm optimization problems.
3. Numerical results are presented in Sect. 4 to compare the proposed approach with other approaches, including the convex relaxation approach and TNN, in terms of recoverability.

## 2 Ky Fan 2- $k$ -norm-based models

The Ky Fan 2- $k$ -norm is the  $\ell_2$ -norm of the vector of  $k$  largest singular values with  $k \leq \min\{m, n\}$ . Thus we have:

$$\|A\|_{k,2} = \left( \sum_{i=1}^k \sigma_i^2(A) \right)^{\frac{1}{2}} \leq \|A\|_F = \left( \sum_{i=1}^{\min\{m,n\}} \sigma_i^2(A) \right)^{\frac{1}{2}},$$

where  $\|\cdot\|_F$  is the Frobenius norm. Now consider the dual Ky Fan 2- $k$ -norm and use the definition of the dual norm, we obtain the following inequality:

$$\|A\|_F^2 = \langle A, A \rangle \leq \|A\|_{k,2}^* \cdot \|A\|_{k,2}.$$

Thus we have:

$$\|A\|_{k,2} \leq \|A\|_F \leq \|A\|_{k,2}^*, \quad k \leq \min\{m, n\}. \quad (4)$$

It is clear that these inequalities become equalities if and only if  $\text{rank}(A) \leq k$ . It shows that to find a low-rank matrix  $X$  that satisfies  $\mathcal{A}(X) = \mathbf{b}$  with  $\text{rank}(X) \leq k$ , we can solve either the following optimization problem

$$\begin{aligned} \min \quad & \frac{\|X\|_{k,2}^*}{\|X\|_F} \\ \text{s.t.} \quad & \mathcal{A}(X) = \mathbf{b}, \end{aligned} \quad (5)$$

or

$$\begin{aligned} \min \quad & \|X\|_{k,2}^* - \|X\|_F \\ \text{s.t.} \quad & \mathcal{A}(X) = \mathbf{b}. \end{aligned} \quad (6)$$

It is straightforward to see that these non-convex optimization problems can be used to recover low-rank matrices as stated in the following theorem given the norm inequalities in (4).

**Theorem 1** *If there exists a matrix  $X \in \mathbb{R}^{m \times n}$  such that  $\text{rank}(X) \leq k$  and  $\mathcal{A}(X) = \mathbf{b}$ , then  $X$  is an optimal solution of (5) and (6).*

Given the result in Theorem 1, the exact recovery of a low-rank matrix using (5) or (6) relies on the uniqueness of the low-rank solution of  $\mathcal{A}(X) = \mathbf{b}$ . Recht et al. [21] generalized the restricted isometry property of vectors introduced by Candès and Tao [5] to matrices and used it to provide sufficient conditions on the uniqueness of these solutions.

**Definition 1** (Recht et al. [21]) For every integer  $k$  with  $1 \leq k \leq \min\{m, n\}$ , the  $k$ -restricted isometry constant is defined as the smallest number  $\delta_k(\mathcal{A})$  such that

$$(1 - \delta_k(\mathcal{A})) \|X\|_F \leq \|\mathcal{A}(X)\|_2 \leq (1 + \delta_k(\mathcal{A})) \|X\|_F \quad (7)$$

holds for all matrices  $X$  of rank at most  $k$ .

Using Theorem 3.2 in Recht et al. [21], we can obtain the following exact recovery result for (5) and (6).

**Theorem 2** Suppose that  $\delta_{2k}(\mathcal{A}) < 1$  and there exists a matrix  $X \in \mathbb{R}^{m \times n}$  which satisfies  $\mathcal{A}(X) = \mathbf{b}$  and  $\text{rank}(X) \leq k$ , then  $X$  is the unique solution to (5) and (6), which implies exact recoverability.

The condition in Theorem 2 is indeed stronger than those obtained for the nuclear norm approach such as the condition of  $\delta_{5k}(\mathcal{A}) < 1/10$  for exact recovery in Recht et al. [21, Theorem 3.3]. The non-convex optimization problems (5) and (6) use norm ratio and difference. When  $k = 1$ , the norm ratio and difference are computed between the nuclear and Frobenius norm. The idea of using these norm ratios and differences with  $k = 1$  has been used to generate non-convex sparse generalizer in the vector case, i.e.,  $m = 1$ . Yin et al. [24] investigated the ratio  $\ell_1/\ell_2$  while Yin et al. [25] analyzed the difference  $\ell_1 - \ell_2$  in compressed sensing. Note that even though optimization formulations based on these norm ratios and differences are non-convex, they are still relaxations of  $\ell_0$ -norm minimization problem unless the sparsity level of the optimal solution is  $s = 1$ . Our proposed approach is similar to the idea of the truncated difference of the nuclear norm and Frobenius norm discussed in Ma et al [16]. Given a parameter  $t \geq 0$ , the truncated difference is defined as

$$\|A\|_{*,t-F} = \sum_{i=t+1}^{\min\{m,n\}} \sigma_i(A) - \left( \sum_{i=t+1}^{\min\{m,n\}} \sigma_i^2(A) \right)^{\frac{1}{2}} \geq 0.$$

For  $t \geq k - 1$ , the problem of truncated difference minimization can be used to recover matrices with rank at most  $k$  given that  $\|X\|_{*,t-F} = 0$  if  $\text{rank}(X) \leq t + 1$ . Similar results for exact recovery as in Theorem 2 are provided in Theorem 3.7(a) in Ma et al [16]. Despite the similarity with respect to the recovery results, the problems (5) and (6) are motivated from a different perspective.

The proposed approach uses the inequality between the dual Ky Fan 2- $k$ -norm and the Frobenius norm. Given the norm inequalities in (4), we can also generate different non-convex optimization problems using the Ky Fan 2- $k$ -norm instead of its dual together with the Frobenius norm. These problems put the Ky Fan 2- $k$ -norm in the non-convex objective term of the objective, which will be represented only via a (linear) approximation during the inner iterations of the DCA framework, which will be discussed in the next section. This is contrast to our method, where the dual Ky Fan 2- $k$  norm appears in the convex term and therefore is exactly represented on outer iterations, which we believe is preferable. Note that if we use 1-norms, i.e., nuclear norm instead of Frobenius norm and Ky Fan 1- $k$ -norm,  $\|\cdot\|_{k,1}$ , instead of 2- $k$ -norm, we get the TNN method directly from the difference model.

We believe that, instead of the (dual) Ky Fan 1- $k$ -norm, it is preferable in general to use the (dual) Ky Fan 2- $k$ -norm given its property regarding rank- $k$  approximation (see further remarks on this matter in [7]). This explains why the models (5) and (6) are proposed. We are now going to discuss how to solve these problems next.

### 3 Numerical algorithm

#### 3.1 Difference of convex functions algorithms

We start with the problem (5). It can be reformulated as

$$\begin{aligned} \max_{\mathbf{Z}, z} \quad & \|\mathbf{Z}\|_F^2 \\ \text{s.t.} \quad & \|\mathbf{Z}\|_{k,2}^* \leq 1, \\ & \mathcal{A}(\mathbf{Z}) - z \cdot \mathbf{b} = \mathbf{0}, \\ & z > 0, \end{aligned} \quad (8)$$

with the change of variables  $z = 1/\|\mathbf{X}\|_{k,2}^*$  and  $\mathbf{Z} = \mathbf{X}/\|\mathbf{X}\|_{k,2}^*$ . The compact formulation is

$$\min_{\mathbf{Z}, z} \delta_{\mathcal{Z}}(\mathbf{Z}, z) - \|\mathbf{Z}\|_F^2/2, \quad (9)$$

where  $\mathcal{Z}$  is the feasible set of the problem (8) and  $\delta_{\mathcal{Z}}(\cdot)$  is the indicator function of  $\mathcal{Z}$ . The problem (9) is a difference of convex functions (d.c.) optimization problem (see, e.g. [19]). The difference of convex functions algorithm DCA proposed in [19] can be applied to the problem (9) as shown in Algorithm 1.

#### Algorithm 1 (DCA-R)

**Step 1.** Start with  $(\mathbf{Z}^0, z^0) = (\mathbf{X}^0/\|\mathbf{X}^0\|_{k,2}^*, 1/\|\mathbf{X}^0\|_{k,2}^*)$  for some  $\mathbf{X}^0$  such that  $\mathcal{A}(\mathbf{X}^0) = \mathbf{b}$  and set  $s = 0$ .

**Step 2.** Update  $(\mathbf{Z}^{s+1}, z^{s+1})$  as an optimal solution of the following convex optimization problem

$$\begin{aligned} \max_{\mathbf{Z}, z} \quad & \langle \mathbf{Z}^s, \mathbf{Z} \rangle \\ \text{s.t.} \quad & \|\mathbf{Z}\|_{k,2}^* \leq 1, \\ & \mathcal{A}(\mathbf{Z}) - z \cdot \mathbf{b} = \mathbf{0}, \\ & z > 0. \end{aligned} \quad (10)$$

**Step 3.** Set  $s \leftarrow s + 1$  and repeat Step 2.

Let  $\mathbf{X}^s = \mathbf{Z}^s/z^s$  and use the general convergence analysis of DCA (see, e.g., Theorem 3.7 in [20]), we can obtain the following convergence results.

**Proposition 1** *Given the sequence  $\{\mathbf{X}^s\}$  obtained from the DCA for the problem (9), the following statements are true.*

- (i) *The sequence  $\left\{ \frac{\|\mathbf{X}^s\|_{k,2}^*}{\|\mathbf{X}^s\|_F} \right\}$  is non-increasing and convergent.*
- (ii)  $\left\| \frac{\mathbf{X}^{s+1}}{\|\mathbf{X}^{s+1}\|_{k,2}^*} - \frac{\mathbf{X}^s}{\|\mathbf{X}^s\|_{k,2}^*} \right\|_F \rightarrow 0$  when  $s \rightarrow \infty$ .

The convergence results show that the DCA improves the objective function value of the ratio minimization problem (5). The DCA can stop if  $(Z^s, z^s) \in \mathcal{O}(Z^s)$ , where  $\mathcal{O}(Z^s)$  is the set of optimal solution of (10) and  $(Z^s, z^s)$  which satisfied this condition is called a critical point. Note that (local) optimal solutions of (9) can be shown to be critical points. The following proposition shows that an equivalent condition for critical points.

**Proposition 2**  $(Z^s, z^s) \in \mathcal{O}(Z^s)$  if and only if  $Y = \mathbf{0}$  is an optimal solution of the following optimization problem

$$\begin{aligned} \min_Y \quad & \|X^s + Y\|_{k,2}^* - \frac{\|X^s\|_{k,2}^*}{\|X^s\|_F^2} \cdot \langle X^s, Y \rangle \\ \text{s.t.} \quad & \mathcal{A}(Y) = \mathbf{0}. \end{aligned} \quad (11)$$

**Proof** Consider  $Y \in \text{Null}(\mathcal{A})$ , i.e.,  $\mathcal{A}(Y) = \mathbf{0}$ , we then have:

$$\left( \frac{X^s + Y}{\|X^s + Y\|_{k,2}^*}, \frac{1}{\|X^s + Y\|_{k,2}^*} \right) \in \mathcal{Z}.$$

We compute the objective function value in (10) for this solution and compare that of  $(Z^s, z^s)$ . We have:  $\langle \frac{X^s}{\|X^s\|_{k,2}^*}, \frac{X^s + Y}{\|X^s + Y\|_{k,2}^*} \rangle \leq \langle \frac{X^s}{\|X^s\|_{k,2}^*}, \frac{X^s}{\|X^s\|_{k,2}^*} \rangle$  is equivalent to

$$\|X^s + Y\|_{k,2}^* - \frac{\|X^s\|_{k,2}^*}{\|X^s\|_F^2} \cdot \langle X^s, Y \rangle \geq \|X^s\|_{k,2}^*.$$

When  $Y = \mathbf{0}$ , we achieve the equality. We have:  $(Z^s, z^s) \in \mathcal{O}(Z^s)$  if and only the above inequality holds for all  $Y \in \text{Null}(\mathcal{A})$ , which means  $f(Y; X^s) \geq f(\mathbf{0}; X^s)$  for all  $Y \in \text{Null}(\mathcal{A})$ , where  $f(Y; X) = \|X + Y\|_{k,2}^* - \frac{\|X\|_{k,2}^*}{\|X\|_F^2} \cdot \langle X, Y \rangle$ . Clearly, it is equivalent to the fact that  $Y = \mathbf{0}$  is an optimal solution of (11).  $\square$

The result of Proposition 2 shows the similarity between the norm ratio minimization problem (5) and the norm different minimization problem (6) with respect to the implementation of the DCA. It is indeed that the problem (6) is a DC optimization problem with the objective function as the difference between the dual Ky Fan 2-k-norm and the Frobenius norm together with a convex feasible set. Given that the linear approximation of the Frobenius norm is  $\|X + Y\|_F \approx \|X\|_F + \frac{1}{\|X\|_F} \langle X, Y \rangle$  for  $X \neq \mathbf{0}$ , the DCA for (6) can be described as Algorithm 2 as follows.

### Algorithm 2 (DCA-D)

**Step 1.** Start with some  $X^0$  such that  $\mathcal{A}(X^0) = \mathbf{b}$  and set  $s = 0$ .

**Step 2.** Update  $X^{s+1} = X^s + Y$ , where  $Y$  is an optimal solution of the following convex optimization problem

$$\begin{aligned} \min_Y \quad & \|X^s + Y\|_{k,2}^* - \frac{1}{\|X^s\|_F} \cdot \langle X^s, Y \rangle \\ \text{s.t.} \quad & \mathcal{A}(Y) = \mathbf{0}. \end{aligned} \quad (12)$$

**Step 3.** Set  $s \leftarrow s + 1$  and repeat Step 2.

Note that in (12), we change the variable from  $X$  to  $Y = X - X^s$ , which results in the constraints  $\mathcal{A}(Y) = \mathbf{0}$  for  $Y$  given that  $\mathcal{A}(X^s) = \mathbf{b}$ . Now, it is clear that  $X^s$  is a critical point

for the problem (6) if and only if  $Y$  is an optimal solution of (12). Both problems (11) and (12) can be written in the general form as

$$\begin{aligned} \min_Y \quad & \|X^s + Y\|_{k,2}^* - \alpha(X^s) \cdot \langle X^s, Y \rangle \\ \text{s.t.} \quad & \mathcal{A}(Y) = \mathbf{0}, \end{aligned} \quad (13)$$

where  $\alpha(X) = \frac{\|X\|_{k,2}^*}{\|X\|_F^2}$  for (11) and  $\alpha(X) = \frac{1}{\|X\|_F}$  for (12), respectively. Given that  $\mathcal{A}(X^s) = \mathbf{b}$ , these problems can be written as

$$\begin{aligned} \min_X \quad & \|X\|_{k,2}^* - \alpha(X^s) \cdot \langle X^s, X \rangle \\ \text{s.t.} \quad & \mathcal{A}(X) = \mathbf{b}. \end{aligned} \quad (14)$$

The following proposition shows that  $X^s$  is a critical point of the problem (14) for many functions  $\alpha(\cdot)$  if  $\text{rank}(X^s) \leq k$ .

**Proposition 3** *If  $\text{rank}(X^s) \leq k$ ,  $X^s$  is a critical point of the problem (14) for any function  $\alpha(\cdot)$  which satisfies*

$$\frac{1}{\|X\|_F} \leq \alpha(X) \leq \frac{\|X\|_{k,2}^*}{\|X\|_F^2}. \quad (15)$$

**Proof** If  $\text{rank}(X^s) \leq k$ , we have:  $\alpha(X^s) = 1/\|X^s\|_{k,2}$  since  $\|X^s\|_{k,2} = \|X^s\|_F = \|X^s\|_{k,2}^*$ . Given that

$$\partial \|A\|_{k,2}^* = \arg \max_{X: \|X\|_{k,2} \leq 1} \langle X, A \rangle,$$

we have:  $\alpha(X^s) \cdot X^s \in \partial \|X^s\|_{k,2}^*$ . Thus for all  $Y$ , the following inequality holds:

$$\|X^s + Y\|_{k,2}^* - \|X^s\|_{k,2}^* \geq \langle \alpha(X^s) \cdot X^s, Y \rangle.$$

It implies  $Y = \mathbf{0}$  is an optimal solution of the problem (13) since the optimality condition is

$$\|X^s + Y\|_{k,2}^* - \|X^s\|_{k,2}^* \geq \langle \alpha(X^s) \cdot X^s, Y \rangle, \quad \forall Y : \mathcal{A}(Y) = \mathbf{0}.$$

Thus  $X^s$  is a critical point of the problem (14).  $\square$

Proposition 3 shows that one can potentially use different functions  $\alpha(\cdot)$  such as  $\alpha(X) = \frac{1}{\|X\|_{k,2}}$  for the sub-problem in the general DCA framework to solve the original problem. We are going to experiment with different pre-determined functions  $\alpha(\cdot)$  in the numerical section. Now, the generalized sub-problem (14) is a convex optimization problem, which can be formulated as a semidefinite optimization problem given the following calculation of the dual Ky Fan 2- $k$ -norm provided in [7]:

$$\begin{aligned} \|X\|_{k,2}^* &= \min p + \text{trace}(\mathbf{R}) \\ \text{s.t.} \quad & kp - \text{trace}(\mathbf{P}) = 0, \\ & p\mathbf{I} - \mathbf{P} \succeq 0, \\ & \begin{pmatrix} \mathbf{P} & -\frac{1}{2}\mathbf{X}^T \\ -\frac{1}{2}\mathbf{X} & \mathbf{R} \end{pmatrix} \succeq 0. \end{aligned} \quad (16)$$

In order to implement the DCA, one also needs to consider how to find the initial solution  $X^0$ . We can use the nuclear norm minimization problem (1), the convex relaxation of the

rank minimization problem, to find  $X^0$ . A similar approach is to use the following dual Ky Fan 2- $k$ -norm minimization problem to find  $X^0$  given its low-rank properties:

$$\begin{aligned} \min_X \quad & \|X\|_{k,2}^* \\ \text{s.t.} \quad & \mathcal{A}(X) = b. \end{aligned} \quad (17)$$

This initial problem can be considered as an instance of (14) with  $X^0 = 0$ . Given that  $\langle X^0, X \rangle = 0$ , one can set any value for  $\alpha(0)$ , e.g.,  $\alpha(0) = 1$ . It is equivalent to starting the iterative algorithm with  $X^0 = 0$  one step ahead. All of these instances of (14) can be solved with a semidefinite optimization solver, which will not be efficient for large problem instances. In the next section, we will develop a primal-dual proximal point algorithm to solve (14).

### 3.2 Proximal point algorithm

Liu et al. [15] proposed a proximal point algorithm (PPA) framework to solve the nuclear norm minimization problem (1). We will develop a similar algorithm to solve the following general problem

$$\begin{aligned} \min_X \quad & f(X) = \|X\|_{k,2}^* + \langle C, X \rangle \\ \text{s.t.} \quad & \mathcal{A}(X) = b, \end{aligned} \quad (18)$$

where  $C \in \mathbb{R}^{m \times n}$  is a parameter, of which (14) is an instance.

This section is divided into four subsections as follows.

1. We first write down the dual problem, its Moreau-Yoshida regularization, and finally a spherical quadratic approximation of the smooth term. Deriving the formulation of norm-plus-spherical is the main task needed in order to use the PPA framework.
2. Next, we argue by symmetry that the norm-plus-spherical subproblem can be rewritten as a vector-norm optimization problem involving only the largest singular values of a certain matrix.
3. This vector-norm optimization problem can be solved in closed form. This requires several technical results whose proofs are deferred to Appendix A.
4. Finally, we summarize all the computations of the PPA method.

#### 3.2.1 Deriving the norm-plus-spherical subproblem

The steps involved in setting up the PPA framework are formulating the dual, its Moreau-Yoshida regularization, and the quadratic approximation thereof. The dual problem can be written simply as  $\max_z g(z)$ , where  $g$  is the concave function defined by

$$g(z) = \min_X f(X) + \langle z, b - \mathcal{A}(X) \rangle. \quad (19)$$

The Moreau-Yoshida regularization of  $g$  can be computed by applying the strong duality (or minimax theory) result in Rockafellar [22]:



$$\begin{aligned}
G_\lambda(\mathbf{z}) &:= \max_{\mathbf{y}} \left\{ g(\mathbf{y}) - \frac{1}{2\lambda} \|\mathbf{z} - \mathbf{y}\|^2 \right\} \\
&= \max_{\mathbf{y}} \min_X \left\{ f(\mathbf{X}) + \langle \mathbf{y}, \mathbf{b} - \mathcal{A}(\mathbf{X}) \rangle - \frac{1}{2\lambda} \|\mathbf{z} - \mathbf{y}\|^2 \right\} \\
&= \min_X \max_{\mathbf{y}} \left\{ f(\mathbf{X}) + \langle \mathbf{y}, \mathbf{b} - \mathcal{A}(\mathbf{X}) \rangle - \frac{1}{2\lambda} \|\mathbf{z} - \mathbf{y}\|^2 \right\}. \quad (20)
\end{aligned}$$

The inner optimization problem can be solved easily with  $\mathbf{y}^* = \mathbf{z} + (\mathbf{b} - \mathcal{A}(\mathbf{X}))$ . Thus we can write down the Moreau-Yoshida regularization of  $g$  as follows:

$$G_\lambda(\mathbf{z}) = -\frac{1}{2\lambda} \|\mathbf{z}\|^2 + \min_X \{f(\mathbf{X}) + \Psi_\lambda(\mathbf{X}; \mathbf{z})\}, \quad (21)$$

where

$$\Psi_\lambda(\mathbf{X}; \mathbf{z}) = \frac{1}{2\lambda} \|\mathbf{z} + \lambda(\mathbf{b} - \mathcal{A}(\mathbf{X}))\|^2. \quad (22)$$

The function  $\Psi_\lambda(\cdot; \mathbf{z})$  is convex and differentiable with the gradient  $\nabla_X \Psi_\lambda(\mathbf{X}; \mathbf{z}) = -\mathcal{A}^*(\mathbf{z} + \lambda(\mathbf{b} - \mathcal{A}(\mathbf{X})))$ , which is Lipschitz continuous with modulus  $L = \lambda \|\mathcal{A}\|_2^2$ . In order to find the proximal point mapping  $p_\lambda$  associated with  $g$ , we need to find an optimal solution  $\mathbf{X}^*$  for the inner minimization problem in (21) since  $p_\lambda(\mathbf{z}) = \mathbf{z} + \lambda(\mathbf{b} - \mathcal{A}(\mathbf{X}^*))$ . For a fixed  $\mathbf{z}$ , the inner minimization problem we need to consider is

$$\min_X \{f(\mathbf{X}) + P(\mathbf{X})\}, \quad (23)$$

where  $P(\mathbf{X}) = \Psi_\lambda(\mathbf{X}; \mathbf{z})$ . As discussed in Liu et al. [15], we will use an accelerated proximal gradient (APG) algorithm for this problem. The APG algorithm in each iteration needs to solve the following approximation of the sum  $f(\mathbf{X}) + P(\mathbf{X})$  at the current solution  $\mathbf{Z}$ :

$$\begin{aligned}
Q_t(\mathbf{X}; \mathbf{Z}) &= f(\mathbf{X}) + P(\mathbf{Z}) + \langle \nabla P(\mathbf{Z}), \mathbf{X} - \mathbf{Z} \rangle + \frac{t}{2} \|\mathbf{X} - \mathbf{Z}\|_F^2 \\
&= f(\mathbf{X}) + \frac{t}{2} \|\mathbf{X} - G_t(\mathbf{Z})\|_F^2 + P(\mathbf{Z}) - \frac{1}{2t} \|\nabla P(\mathbf{Z})\|^2,
\end{aligned}$$

where  $G_t(\mathbf{Z}) = \mathbf{Z} - \frac{1}{t} \nabla P(\mathbf{Z})$  and  $t > 0$  is a parameter.

This problem has a unique solution  $S_t(\mathbf{Z})$  given the strong convexity of the function. We have:

$$f(\mathbf{X}) + \frac{t}{2} \|\mathbf{X} - G_t(\mathbf{Z})\|_F^2 = \|\mathbf{X}\|_{k,2}^* + \langle \mathbf{C}, \mathbf{X} \rangle + \frac{t}{2} \|\mathbf{X} - G_t(\mathbf{Z})\|_F^2,$$

which means  $S_t(\mathbf{Z})$  is the minimizer of the problem:

$$\min_X \left\{ \|\mathbf{X}\|_{k,2}^* + \frac{t}{2} \left\| \mathbf{X} - \left( G_t(\mathbf{Z}) - \frac{1}{t} \mathbf{C} \right) \right\|_F^2 \right\},$$

which is an instance of the following optimization problem

$$\min_X \|\mathbf{X}\|_{k,2}^* + \frac{1}{2\lambda} \|\mathbf{X} - \mathbf{Y}\|_F^2, \quad (24)$$

where  $\lambda = 1/t > 0$  and  $\mathbf{Y} = G_t(\mathbf{Z}) - \frac{1}{t} \mathbf{C}$  are parameters. For the nuclear norm minimization problem, the dual Ky Fan 2- $k$ -norm is replaced by the nuclear norm and the resulting problem

can be solved easily with the SVD decomposition and the shrinkage operator (see, e.g., Liu et al. [15] and references therein).

### 3.2.2 Reduction of the norm-plus-spherical problem to a convex vector-norm problem

Using the orthogonal invariance of  $\|\cdot\|_{k,2}^*$ , we first prove that the optimal solution  $\mathbf{X}^*$  of (24) can be constructed from the optimal solution  $\mathbf{x}^*$  of the following optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{k,2}^* + \frac{1}{2\lambda} \|\mathbf{x} - \mathbf{y}\|_2^2, \quad (25)$$

where  $\|\mathbf{x}\|_{k,2}^*$  is the dual of the Ky Fan 2- $k$ -vector norm of  $\mathbf{x}$ ,  $\|\mathbf{x}\|_{k,2} = \left( \sum_{i=1}^k |x|_{(i)}^2 \right)^{\frac{1}{2}}$ , with  $|x|_{(i)}$  is the  $(n-i+1)$ -th order statistic of  $|\mathbf{x}|$ .  $\|\cdot\|_{k,2}^*$  is the corresponding symmetric gauge function of  $\|\cdot\|_{k,2}$  and the above result which we shall prove can actually be applied to any orthogonally invariant matrix norm  $\|\cdot\|$  and its corresponding symmetric gauge function  $\|\cdot\|$ . We start with the following lemma.

**Lemma 1** Consider the optimal solution  $\mathbf{x}^*$  of the following optimization problem:

$$\min_{\mathbf{x}} \|\mathbf{x}\| + \frac{1}{2\lambda} \|\mathbf{x} - \mathbf{y}\|_2^2, \quad (26)$$

where  $\|\cdot\|$  is a symmetric gauge function. The following statements are true.

- (i) If  $\mathbf{y} \geq \mathbf{0}$  then  $\mathbf{x}^* \geq \mathbf{0}$ .
- (ii) If  $y_i \geq y_j$  then  $x_i^* \geq x_j^*$ .

**Proof** (i) If  $\mathbf{x}^* \not\geq \mathbf{0}$ , without loss of generality, let us assume that  $x_1^* < 0$ . Construct the solution  $\mathbf{x}$  from  $\mathbf{x}^*$  by inverting the sign of the first element,  $x_1 = -x_1^* > 0$ . We have:  $\|\mathbf{x}\| = \|\mathbf{x}^*\|$  and  $\|\mathbf{x} - \mathbf{y}\|_2 \leq \|\mathbf{x}^* - \mathbf{y}\|_2$ . Thus  $\mathbf{x} \neq \mathbf{x}^*$  is also an optimal solution of (26), which contradicts the uniqueness of the optimal solution  $\mathbf{x}^*$ . Thus we have  $\mathbf{x}^* \geq \mathbf{0}$ .

(ii) Assume that  $x_i^* < x_j^*$ , we construct another solution  $\mathbf{x}$  as follows.

$$x_k = x_k^*, k \neq i, j, \quad x_i = x_j^*, x_j = x_i^*.$$

We have:  $\|\mathbf{x}\| = \|\mathbf{x}^*\|$  since  $\|\cdot\|$  is a symmetric gauge function. We have:

$$\begin{aligned} \|\mathbf{x}^* - \mathbf{y}\|_2^2 - \|\mathbf{x} - \mathbf{y}\|_2^2 &= (x_i^* - y_i)^2 + (x_j^* - y_j)^2 - (x_i^* - y_j)^2 - (x_j^* - y_i)^2 \\ &= (x_i^* - x_j^*)(y_j - y_i) \geq 0. \end{aligned}$$

Thus  $\mathbf{x} \neq \mathbf{x}^*$  is also an optimal solution of (26), which contradicts the uniqueness of the optimal solution  $\mathbf{x}^*$ . Thus we have  $x_i^* \geq x_j^*$ . □

We can now state the main result for the optimization problem with the general orthogonally invariant norm  $\|\cdot\|$ ,

$$\min_{\mathbf{X}} \left\{ \|\mathbf{X}\| + \frac{1}{2\lambda} \|\mathbf{X} - \mathbf{Y}\|_F^2 \right\}. \quad (27)$$

**Proposition 4** If  $Y = U \text{Diag}(\sigma(Y)) V^T$  is a singular decomposition of  $Y$ , then the optimal solution  $X^*$  of (27) can be calculated as

$$X^* = U \text{Diag}(x^*) V^T,$$

where  $x^*$  is the optimal solution of (26) with  $y = \sigma(Y)$ .

**Proof** Problem (27) with the general  $\|\cdot\|$  has a unique optimal solution  $X^*$  that satisfies the following optimality condition:

$$0 \in \partial \|\|X^*\| + \frac{1}{\lambda} (X^* - Y) \Leftrightarrow Y \in \lambda \partial \|\|X^*\| + X^*.$$

Similarly, the optimal solution  $x^*$  satisfies the condition

$$0 \in \partial \|x^*\| + \frac{1}{\lambda} (x^* - y) \Leftrightarrow y \in \lambda \partial \|x^*\| + x^*.$$

We have:  $y = \sigma(Y) \geq 0$ . Thus according to Lemma 1,  $x^* \geq 0$ . Similarly,  $y_1 \geq \dots \geq y_n$  implies that  $x_1^* \geq \dots \geq x_n^*$ .

Now let us consider  $X = U \text{Diag}(x^*) V^T$ , we have:  $\sigma(X) = x^*$ . According to Ziętak [26], the subgradient of  $\|\cdot\|$  is

$$\partial \|\|X\| = \left\{ Z : Z = \bar{U} \text{Diag}(\tau) \bar{V}^T, \tau \in \partial \|\sigma(X)\| \right\},$$

where  $X = \bar{U} \text{Diag}(\sigma(X)) \bar{V}^T$  is any singular value decomposition of  $X$ . We have:  $\frac{1}{\lambda} (y - x^*) \in \partial \|\sigma(X)\|$ , thus

$$Z = \frac{1}{\lambda} U \text{Diag}(y - x^*) V^T \in \partial \|\|X\| \Leftrightarrow Y \in \lambda \partial \|\|X\| + X.$$

$X$  satisfies the optimality condition, thus  $U \text{Diag}(x^*) V^T$  is the (unique) optimal solution of (27).  $\square$

### 3.2.3 Solution of the vector-norm problem

Proposition 4 shows that we can focus on (25). The following proposition is proved in Appendix A.

**Proposition 5** The dual Ky Fan 2- $k$ -vector norm can be computed using the following formulation:

$$\|y\|_{k,2}^* = \sqrt{\sum_{j=1}^{k-i^*-1} (|y|_{(j)})^2 + \frac{1}{i^*+1} \left( \sum_{j=k-i^*}^n |y|_{(j)} \right)^2}, \quad (28)$$

where

$$i^* = \max \left\{ i : |y|_{(k-i)} < \frac{1}{i} \left( \sum_{j=k-i+1}^n |y|_{(j)} \right), 0 \leq i \leq k-1 \right\}. \quad (29)$$

We are now ready to solve the problem (25). The optimality condition for (25) can be written as follows:

$$\mathbf{0} \in \partial \|\mathbf{x}\|_{k,2}^* + \frac{1}{\lambda}(\mathbf{x} - \mathbf{y}) \Leftrightarrow \mathbf{y} \in \lambda \partial \|\mathbf{x}\|_{k,2}^* + \mathbf{x}.$$

We shall focus on the case  $\mathbf{y} = \sigma(\mathbf{Y}) \geq \mathbf{0}$  and assume that  $y_1 \geq \dots \geq y_n \geq 0$ . Using the formulation of  $\|\cdot\|_{k,2}^*$  stated in Proposition 5, we can compute the subgradient  $\partial \|\cdot\|_{k,2}^*$  as follows. If  $\mathbf{y} = \mathbf{0}$ , we have:

$$\partial \|\mathbf{y}\|_{k,2}^* = \{\mathbf{z} \in \mathbb{R}^n : \|\mathbf{z}\|_{k,2} \leq 1\}.$$

If  $\mathbf{y} \neq \mathbf{0}$ , compute  $i^* \in [0, k-1]$  using (29). For  $1 \leq j \leq k - i^* - 1$ , we have:

$$\frac{\partial \|\mathbf{y}\|_{k,2}^*}{\partial y_j} = \frac{y_j}{\|\mathbf{y}\|_{k,2}^*}.$$

Let assume that there are  $n_0$  zeros among the last  $n - k + i^* + 1$  elements of  $\mathbf{y}$  starting from  $(k - i^*)$ -th element. For  $k - i^* \leq j \leq n - n_0$ , we have:

$$\frac{\partial \|\mathbf{y}\|_{k,2}^*}{\partial y_j} = \frac{1}{(i^* + 1) \|\mathbf{y}\|_{k,2}^*} \left( \sum_{l=k-i^*}^n y_l \right).$$

Finally we have:

$$(\partial \|\mathbf{y}\|_{k,2}^*)_{[n-n_0+1:n]} = \left\{ \frac{1}{(i^* + 1) \|\mathbf{y}\|_{k,2}^*} \left( \sum_{l=k-i^*}^n y_l \right) \mathbf{z} \in \mathbb{R}^{n_0} : -\mathbf{e}_{n_0} \leq \mathbf{z} \leq \mathbf{e}_{n_0} \right\},$$

where  $\mathbf{e}_{n_0} \in \mathbb{R}^{n_0}$  is the vector of all ones.

Using the formulation of the subgradient at  $\mathbf{x} = \mathbf{0}$  and the uniqueness of the optimal solution, we can show that if  $\|\mathbf{y}\|_{k,2} \leq \lambda$ , then  $\mathbf{x} = \mathbf{0}$  is the optimal solution of (25). Now consider the case  $\|\mathbf{y}\|_{k,2} > \lambda$ . According to Lemma 1, the optimal solution of (25) is nonnegative, i.e.,  $\mathbf{x} \geq \mathbf{0}$ . Assuming that the values of  $i^*$  and  $n_0$  have been given, the optimality conditions can then be written as follows:

$$\mathbf{y}_{[1:k-i^*-1]} = \frac{\alpha + 1}{\alpha} \mathbf{x}_{[1:k-i^*-1]}, \quad (30)$$

$$\mathbf{y}_{[k-i^*:n-n_0]} = \left( \mathbf{I}_{m_0} + \frac{1}{(i^* + 1)\alpha} \mathbf{e}_{m_0} \mathbf{e}_{m_0}^T \right) \mathbf{x}_{[k-i^*:n-n_0]}, \quad (31)$$

$$\mathbf{y}_{[n-n_0+1:n]} \leq \frac{\mathbf{e}_{m_0}^T \mathbf{x}_{[k-i^*:n-n_0]}}{(i^* + 1)\alpha} \mathbf{e}_{n_0}, \quad (32)$$

where  $\alpha = \|\mathbf{x}\|_{k,2}^*/\lambda$ ,  $m_0 = n - n_0 - k + i^* + 1$  and  $\mathbf{I}_{m_0} \in \mathbb{R}^{m_0 \times m_0}$  is the identity matrix.

Using the Sherman-Morrison formula for the inverse of  $\mathbf{I}_{m_0} + \frac{1}{(i^* + 1)\alpha} \mathbf{e}_{m_0} \mathbf{e}_{m_0}^T$ , we obtain the analytic formulation for  $\mathbf{x}$ :

$$\mathbf{x}_{[1:k-i^*-1]} = \frac{\alpha}{\alpha + 1} \mathbf{y}_{[1:k-i^*-1]}, \quad (33)$$

$$\mathbf{x}_{[k-i^*:n-n_0]} = \left( \mathbf{I}_{m_0} - \frac{1}{(i^* + 1)\alpha + m_0} \mathbf{e}_{m_0} \mathbf{e}_{m_0}^T \right) \mathbf{y}_{[k-i^*:n-n_0]}, \quad (34)$$

$$\mathbf{x}_{[n-n_0+1:n]} = \mathbf{0}. \quad (35)$$

The unknown  $\alpha > 0$  satisfies the following equation:

$$\alpha = \frac{1}{\lambda} \sqrt{\frac{\alpha^2}{(\alpha + 1)^2} \|\mathbf{y}_{[1:k-i^*-1]}\|_2^2 + \frac{(i^* + 1)\alpha^2}{[(i^* + 1)\alpha + m_0]^2} \|\mathbf{y}_{[k-i^*:n-n_0]}\|_1^2},$$

or equivalently

$$f(\alpha) = \frac{1}{\lambda} \sqrt{\frac{1}{(\alpha + 1)^2} \|\mathbf{y}_{[1:k-i^*-1]}\|_2^2 + \frac{i^* + 1}{[(i^* + 1)\alpha + m_0]^2} \|\mathbf{y}_{[k-i^*:n-n_0]}\|_1^2} = 1. \quad (36)$$

Given the relationship between  $\mathbf{x}$  and  $\mathbf{y}$ , we have:

$$0 < \alpha = \frac{\|\mathbf{x}\|_{k,2}^*}{\lambda} < \alpha_{\max} = \frac{1}{\lambda} \sqrt{\|\mathbf{y}_{[1:k-i^*-1]}\|_2^2 + \frac{1}{i^* + 1} \|\mathbf{y}_{[k-i^*:n-n_0]}\|_1^2}.$$

The function  $f(\alpha)$  is a (strictly) decreasing function on  $[0, +\infty)$ ; therefore, the necessary conditions are  $f(0) > 1$  and  $f(\alpha_{\max}) < 1$ . If  $f(0) > 1$  and  $f(\alpha_{\max}) < 1$ , the value of  $\alpha$  can be found using the efficient binary search. Given the values of  $i^*$ ,  $n_0$ , and  $\alpha$ , the solution  $\mathbf{x}$  can be computed completely. Note that the optimal solution  $\mathbf{x}$  is unique, which guarantees that if all optimality conditions are satisfied and  $i^*$  is correct according to (29), we indeed obtain the unique optimal solution  $\mathbf{x}$ .

### 3.2.4 PPA summary

The analyses in the preceding paragraphs result in Algorithm 3 that can be used to solve the problem (25).

#### Algorithm 3 (PROX)

**Step 1.** Compute  $\|\mathbf{y}\|_{k,2}^*$ . If  $\|\mathbf{y}\|_{k,2}^* \leq \lambda$ , return  $\mathbf{x} = \mathbf{0}$ . Otherwise, continue.

**Step 2.** Iterate  $n_0 = 1 : n$  and  $i^* = \max\{0, k - 1 + n_0 - n\} : k - 1$ :

**Step 2.1.** Compute  $f(0)$  and  $f(\alpha_{\max})$ . If  $f(0) > 1$  and  $f(\alpha_{\max}) < 1$ , compute  $\alpha$  using (36) and go to Step 2.2. Otherwise, continue to the next iteration.

**Step 2.2.** Compute  $\mathbf{x}$  using (33), (34), and (35). If  $\mathbf{x} \geq \mathbf{0}$ , (32) and (29) are satisfied, return  $\mathbf{x}$ . Otherwise, continue to the next iteration.

Algorithm 3 together with Proposition 4 allows us to solve the problem (24) using the SVD decomposition, which means that we can compute  $S_t(\mathbf{Z})$  efficiently. The APG algorithm for solving (23) can then be described as follows. Given  $\tau_0 = \tau_{-1} = 1$  and  $\mathbf{X}^0 = \mathbf{X}^{-1}$ , each iteration includes the following steps

**Step 1.** Calculate  $\mathbf{Y}^s = \mathbf{X}^s + \frac{\tau_{s-1} - 1}{\tau_s} (\mathbf{X}^s - \mathbf{X}^{s-1})$ .

**Step 2.** Update  $\mathbf{X}^{s+1} = S_{t_s}(\mathbf{Y}^s)$  using Algorithm 3 and Proposition 4.

**Step 3.** Update  $\tau_{s+1} = \frac{1}{2} \left( \sqrt{1 + 4\tau_s^2} + 1 \right)$ .

The additional parameter  $\tau_s$  and how it is updated in each step are necessary for the convergence analysis of the APG algorithm when  $t_s$  is set to be  $L$ , the Lipschitz modulus of the gradient  $\nabla P(\mathbf{Y})$  (see, e.g., Beck and Teboulle [2]). We now can write down the proximal point algorithm (PPA) to solve the problem (18) as follows:

**Algorithm 4 (PPA)**

Given a tolerance  $\varepsilon > 0$ . Input  $\mathbf{z}^0$  and  $\lambda_0 > 0$ . Set  $s = 0$ . Iterate:

**Step 1.** Find an approximate minimizer

$$\mathbf{X}^s \approx \arg \min_{\mathbf{X}} \left\{ F(\mathbf{X}) := \|\mathbf{X}\|_{k,2}^* + \langle \mathbf{C}, \mathbf{X} \rangle + \Psi_{\lambda_s}(\mathbf{X}; \mathbf{z}^s) \right\}, \quad (37)$$

where  $\Psi_{\lambda_s}(\mathbf{X}; \mathbf{z}^s)$  is defined as in (22).

**Step 2.** Compute

$$\mathbf{z}^{s+1} = \mathbf{z}^s + \lambda_s(\mathbf{b} - \mathcal{A}(\mathbf{X}^s)). \quad (38)$$

**Step 3.** If  $\|(\mathbf{z}^{s+1} - \mathbf{z}^s)/\lambda_s\| \leq \varepsilon$ ; stop; else; update  $\lambda_s$ ; end.

The proximal parameter  $\lambda_s$  affects the convergence speeds of both inner optimization problem (37) and the whole PPA algorithm. Similar to the discussion in Doan et al. [6], one can initially fix  $\lambda_0$  and increase  $\lambda_s$  only when the outer algorithm converges too slowly as compared to the inner algorithm. One can use the relative convergence measure  $rel_o^s = \frac{\|\mathbf{b} - \mathcal{A}(\mathbf{X}^s)\|}{\|\mathbf{b}\|}$  for the outer algorithm and  $rel_i^s = \frac{\|\mathbf{X}^s - \mathbf{X}^{s-1}\|_F}{\|\mathbf{X}^s\|_F}$  for the inner algorithm in deciding when to increase  $\lambda_s$ . This PPA algorithm can then be used in Step 2 of the proposed DCA to solve the problem (10), which is an instance of (18). We are now ready to provide numerical results to demonstrate the effectiveness of the proposed approach.

## 4 Numerical results

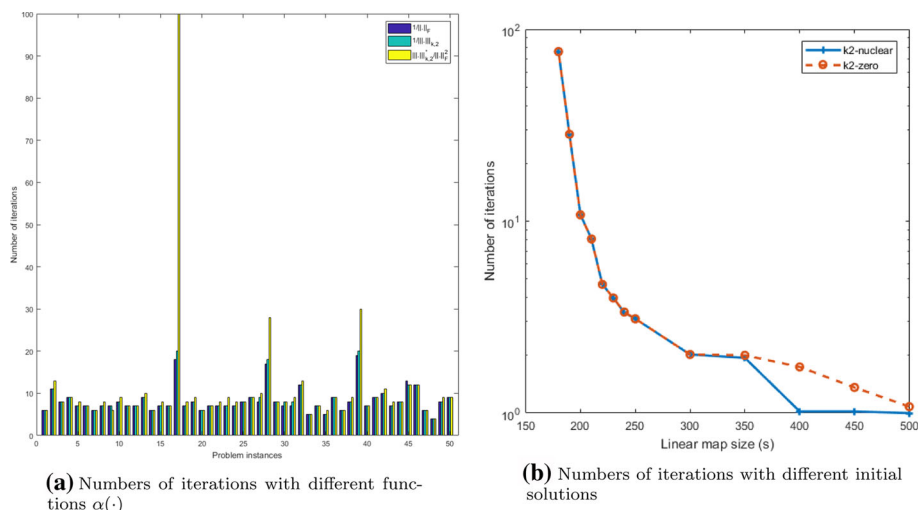
In this section, we are going to discuss the proposed general DCA framework both in terms of computation and recoverability. Similar to Candès and Recht [4], we construct the following the experiment. We generate  $\mathbf{M}$ , an  $m \times n$  matrix of rank  $r$ , by sampling two  $m \times r$  and  $n \times r$  factors  $\mathbf{M}_L$  and  $\mathbf{M}_R$  with i.i.d. Gaussian entries and setting  $\mathbf{M} = \mathbf{M}_L \mathbf{M}_R$ . The linear map  $\mathcal{A}$  is constructed with  $s$  independent Gaussian matrices  $\mathbf{A}_i$  whose entries follows  $\mathcal{N}(0, 1/s)$ , i.e.,

$$\mathcal{A}(\mathbf{X}) = \mathbf{b} \Leftrightarrow \langle \mathbf{A}_i, \mathbf{X} \rangle = \langle \mathbf{A}_i, \mathbf{M} \rangle = b_i, \quad i = 1, \dots, s.$$

We now start with computational settings for the proposed DCA framework.

### 4.1 Computational settings

The proposed general DCA framework allows us to choose the function  $\alpha(\cdot)$  for the sub-problem as well as the initial solution. In order to test different functions  $\alpha(\cdot)$ , we generate  $K = 50$  matrix  $\mathbf{M}$  with  $m = 50$ ,  $n = 40$ , and  $r = 2$ . The dimension of these matrices is  $d_r = r(m + n - r) = 176$ . For each  $\mathbf{M}$ , we generate  $s$  matrices for the random linear map with  $s = 200$ . We set the maximum number of iterations of the algorithm to be  $N_{\max} = 100$  together with the tolerance of  $\epsilon = 10^{-6}$  for  $\|\mathbf{X}^{s+1} - \mathbf{X}^s\|_F$ . For this experiment, the instances are solved using SDPT3 solver [23] for semi-definite optimization problems in Matlab with CVX [9]. The computer used for these numerical experiments is a 64-bit Windows 10 machine with 3.70GHz quad-core CPU, and 32GB RAM. We run the proposed DCA with  $k = r = 2$  and three different pre-determined functions,  $\alpha_1(\mathbf{X}) = \frac{1}{\|\mathbf{X}\|_F}$ ,



**Fig. 1** Computational performance of the DCA under different settings

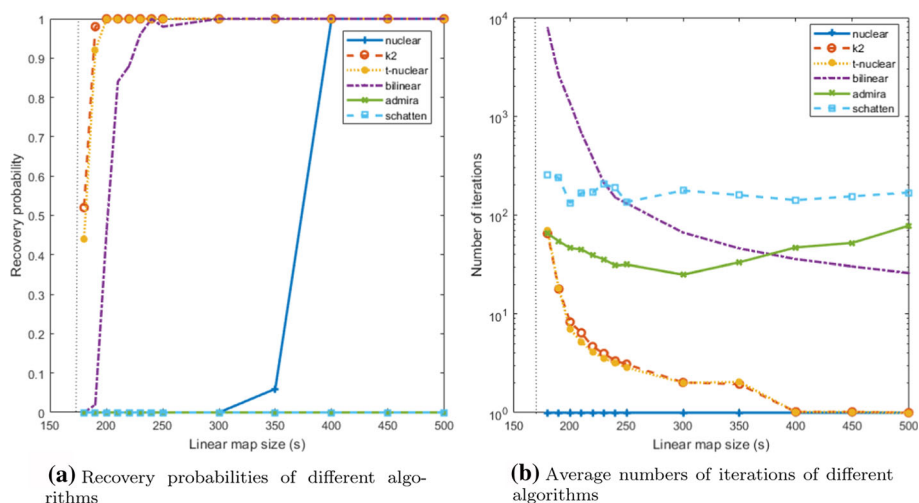
$\alpha_2(X) = \frac{1}{\|X\|_{k,2}}$ , and  $\alpha_3(X) = \frac{\|X\|_{k,2}^*}{\|X\|_F^2}$ . We use the solution obtained from the nuclear optimization formulation (1) as the initial solution for the proposed DCA in this experiment.

Figure 1a shows that all three functions  $\alpha(\cdot)$  perform similarly for most of the instances. However, for difficult instances which require a large number of iterations to converge,  $\alpha_3(\cdot)$  performs the worst as compared to the other two. There is one instance with which the DCA with  $\alpha_3(\cdot)$  does not converge after the maximum number of iterations. Given that  $\alpha_1(\cdot)$  performs slightly better than  $\alpha_2(\cdot)$  in all tested instances, we are going to use  $\alpha_1(\cdot)$  in subsequent experiments.

The next experiment is used to test the options of initial solutions. In this experiment, we vary  $s$  from 180 to 500 and generate  $K = 50$  instances for each value of  $s$ . We run the two variants of the proposed PCA, k2-nuclear with initial solution obtained from (1) and k2-zero with initial solution  $X^0 = \mathbf{0}$  as discussed in Sect. 3.1. Figure 1b shows that two variants perform similarly in terms of average number of iterations for most sizes of the random linear map. k2-nuclear is better when the size of the linear map is large, which can be explained by the recoverability of the convex relaxation approach nuclear using the nuclear optimization formulation (1). We shall therefore use k2-nuclear in comparison with other approaches including nuclear in the next section.

## 4.2 Recoverability

In order to compare the proposed approach with others, we again generate  $K = 50$  random instances with  $m = 50$ ,  $n = 40$ , and  $r = 2$  for different sizes of the linear map ranging from 180 to 500. In order to check whether the matrix  $M$  is recovered, we use the relative error  $\frac{\|X - M\|_F}{\|M\|_F}$  as the performance measure. The matrix  $M$  is considered to be recovered successfully if  $\frac{\|X - M\|_F}{\|M\|_F} \leq \epsilon_t$ , where the threshold  $\epsilon_t$  is set to be  $10^{-6}$  for these experiments. We compare our proposed approach k2-nuclear with nuclear and t-nuclear, an



**Fig. 2** Recovery performance of different algorithms given different sizes of the linear map

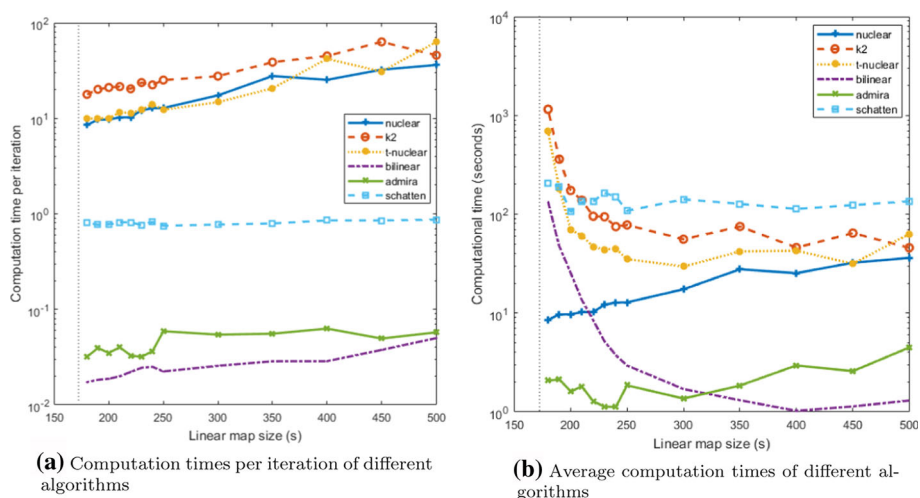
iterative algorithm proposed by Hu et al. [10] for the TNN method. Based on survey paper by Nguyen et al. [18], we also consider the alternating optimization approach proposed by Jain et al. [13] which is used to solve the following equivalent bilinear non-convex optimization problem:

$$\min_{U \in \mathbb{R}^{m \times k}, V \in \mathbb{R}^{n \times k}} \left\| \mathcal{A}(UV^T) - \mathbf{b} \right\|_F^2. \quad (39)$$

Another approach that we consider is the greedy heuristic of atomic decomposition for minimum rank approximation (admira) proposed by Lee and Bresler [14], which aims to find  $k$  rank-one matrices representing the original matrix. Finally, in addition to nuclear, we consider schatten approach proposed by Mohan and Fazel [17] to solve the Schatten  $p$ -norm minimization problem with  $p = 1/2$ , one of the non-convex relaxation methods for which  $k$  is *a priori* unknown (see, e.g., Hu et al. [11]). The tolerance is again set at  $\epsilon = 10^{-6}$  for  $\|X^{s+1} - X^s\|_F$  and the maximum number of iterations is set higher at  $\bar{N}_{\max} = 10000$  for these methods given that their (least-squares) subproblems require less computational efforts.

Figure 2a shows recovery probabilities given different sizes of the linear map. The results show that the proposed algorithm can recover exactly the matrix  $\mathbf{M}$  with 100% rate when  $s \geq 200$  while the nuclear norm approach cannot recover any matrix at all, i.e., 0% rate, if  $s \leq 300$ . Both admira and schatten approaches cannot recover any matrix even though both of them converges with not very large number of iterations as shown in Fig. 2b. This result indicates that the greedy heuristic does not work well for exact recovery. Note that the schatten approach does not solve the non-convex Schatten  $p$ -norm optimization problem exactly, which might explain why it also does not perform well. The bilinear approach is better than the nuclear norm approach with 100% recovery rate when  $s \geq 300$ . However, for small  $s$ , its recovery probability drops significantly to 2% for  $s = 190$  and 0% for  $s = 180$ . The t-nuclear performs comparably with our proposed approach with 100% rate for  $s \geq 200$ . For  $s = 190$  and  $s = 180$ , our approach achieves better recovery rate of 98% and 52% respectively, as compared those of 92% and 44% for the t-nuclear approach.

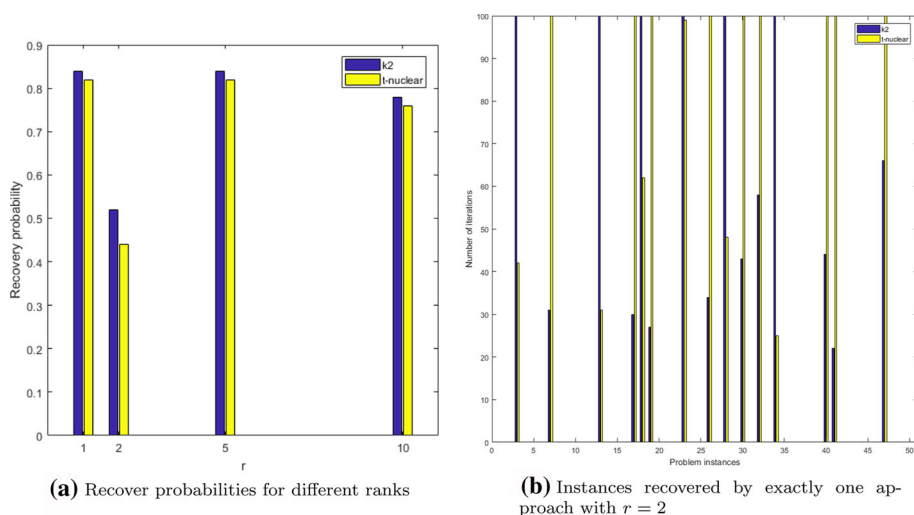




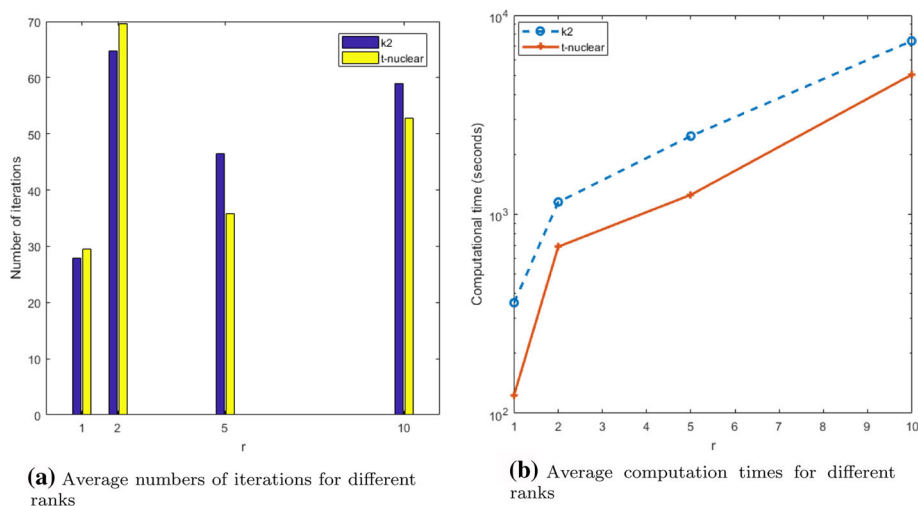
**Fig. 3** Computational performance of different algorithms given different linear map sizes

For all these three approaches, the average number of iterations increases significantly, especially for the bilinear approach, when the size of the linear map decreases. Both our approach and t-nuclear have similar average numbers of iterations. It is interesting to see that we only need 2 extra iterations when  $s = 250$  or 1 extra iteration on average when  $s = 300$  to obtain 100% recover rate when the nuclear norm optimization approach still cannot recover any of the matrices (0% rate). The average numbers of iterations for bilinear are much higher, which implies that it has a much slower convergence rate. On the other hand, the computation time per iteration for bilinear is much lower than those for nuclear, t-nuclear, and our proposed approach. It can be explained by the fact that one needs to solve matrix norm minimization problems using SDPT3 in each iteration of these approaches why the solutions of least-squares subproblems of bilinear can be computed directly in Matlab. The computation times per iteration are shown in Fig. 3a while the average computation times are shown in Fig. 3b. It shows that bilinear is much more efficient when the size of the linear map is large. On the other hand, its average computation time increases significantly when the size of the linear map decreases. As compared with t-nuclear, the average computation time of our approach is higher even though the numbers of iterations are similar. One explanation might be that the dual Ky Fan 2- $k$ -norm optimization problem is more difficult to solve than the nuclear norm optimization problem with the SDPT3 solver. We are going to test the proposed proximal point algorithm that can be used to solve the dual Ky Fan 2- $k$ -norm optimization problem (14) for our proposed approach in the next section.

To conclude this section, we provide some additional recoverability tests with different rank settings including  $r = 1$ ,  $r = 5$  and  $r = 10$ . Note that when  $k = r = 1$ , the dual Ky Fan 2- $k$ -norm is the nuclear norm and the objective function in (6) is  $\|X\|_* - \|X\|_F$  while that of the TNN method is  $\|X\|_* - \|X\|$ , whose sub-problems are both nuclear norm optimization problems. We run the experiments with the following settings of  $(r, s)$ : (1, 95), (2, 180), (5, 440), and (10, 820), in which the size of the linear map is quite close to the dimension  $d_r$ . For all ranks, bilinear approach does not recover any matrix with this size of the linear map and the average number of iterations is very close to the maximum number of  $\tilde{N}_{\max} = 10000$ . Figure 4a shows the recovery probabilities of our proposed approach



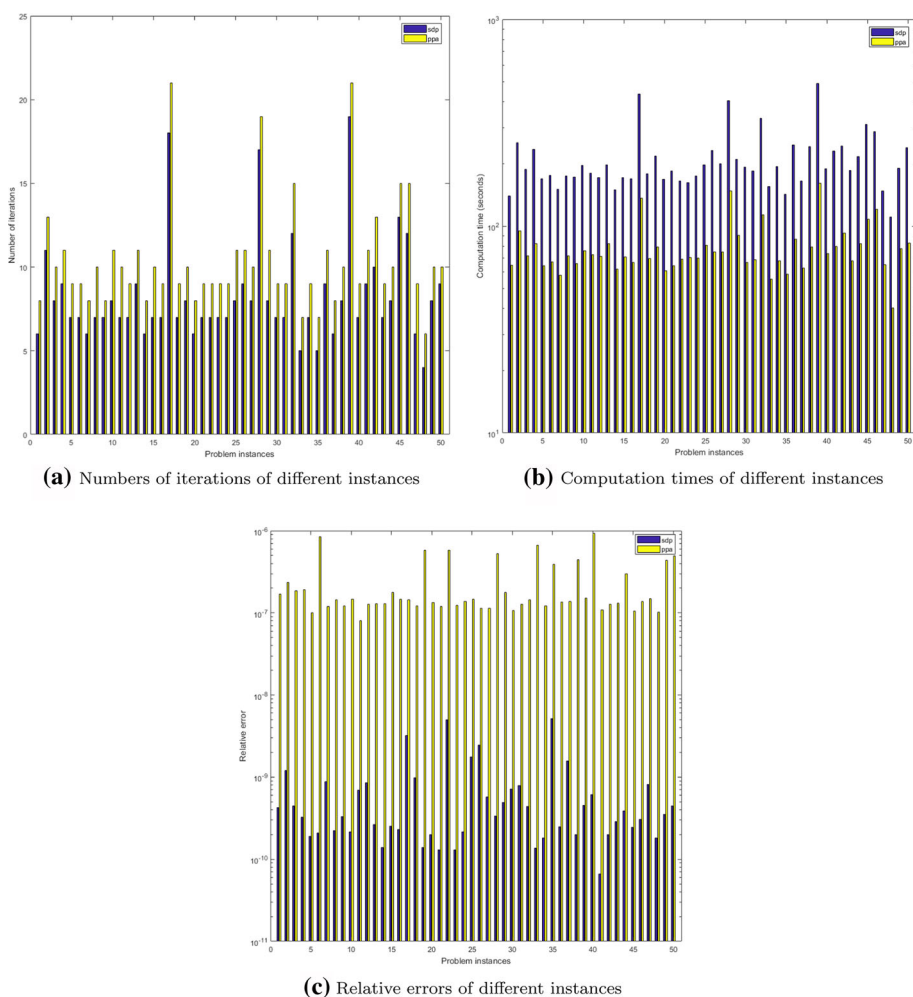
**Fig. 4** Recovery performance of k2 and t-nuclear for different ranks



**Fig. 5** Computational performance of k2 and t-nuclear for different ranks

and t-nuclear. The results indicate that our proposed approach is slightly better than t-nuclear in general in terms of recovery with these small sizes of the linear map. It is interesting to see that there are instances with which one approach can recover the original matrix exactly while the other cannot. Figure 4b shows those instances for  $r = 2$  (maximum number of iterations in general indicates no recovery and vice versa under these settings).

Figure 5a shows that the average numbers of iterations of our proposed approach are slightly higher than those of t-nuclear. The average computation times of our proposed approach are also higher as shown in Fig. 5b. When  $r = 1$ , the average computation time per iteration is still higher even though both sub-problems are nuclear norm minimization problems (10.54 seconds vs. 4.49 seconds). It might be explained by the fact that the sub-

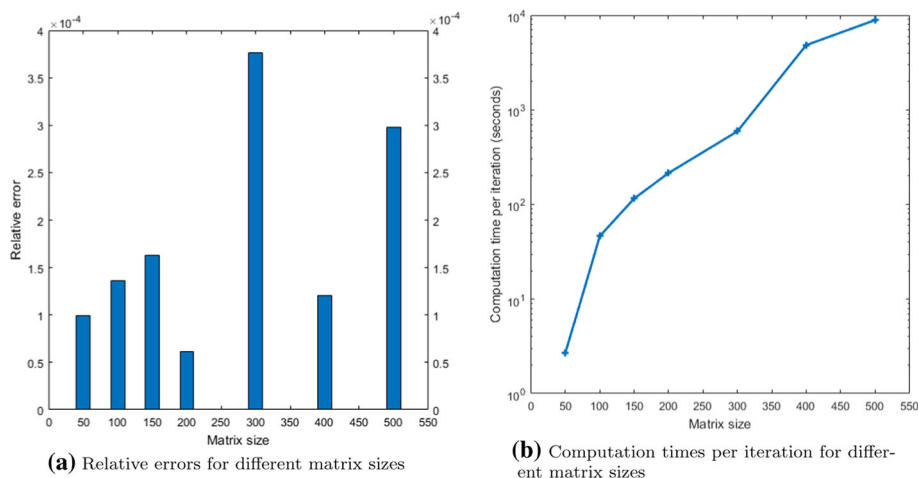


**Fig. 6** Computational performance of k2-zero (sdp) and k2-zero (ppa)

problem in our approach uses the general dual Ky Fan 2- $k$ -norm for  $k = r = 1$  whereas that in t-nuclear uses the function `norm_nuc` in CVX. In the next section, we are going to test whether the proposed proximal point algorithm can be used to solve subproblems in our proposed approach more efficiently.

### 4.3 Proximal point algorithm

In this section, we are going to test the proposed proximal point algorithm used to solve the subproblem (14). We start with the same initial setting of  $m = 50$ ,  $n = 40$ , and  $r = 2$  for the original matrix  $M$  and  $s = 200$  for the random linear map. Instead of k2-nuclear, we run the proposed algorithm k2-zero with initial solution  $X^0 = \mathbf{0}$ , which is easier to set without the need of solving the nuclear norm optimization problem. We use the proximal point algorithm (ppa) for the sub-problems and compare it with the original algorithm which uses the



**Fig. 7** Computational performance of k2-zero (ppa) for different matrix sizes

interior-point method (sdp) for the semidefinite optimization formulation of the subproblem (14) for  $K = 50$  instances. The tolerance is again set to be  $\epsilon = 10^{-6}$  and the maximum number of iterations is  $N_{\max} = 100$ . The tolerance for the proximal point algorithm is set to be  $\epsilon_p = 10^{-8}$  with the maximum number of iterations of  $N_{\max}^p = 100$ . We set  $\lambda_0 = 10$  and will increase  $\lambda_{s+1} = 1.25 \cdot \lambda_s$  if  $rel_o^s > 5 \cdot rel_i^s$ . Figure 6a shows that in general the algorithm k2-zero (ppa) requires more iterations to converge than k2-zero (sdp). In terms of computation time, the algorithm k2-zero (ppa) is definitely better than k2-zero (sdp) in all instances. Figure 6b shows the average computation times for both variants of the proposed algorithm. In terms of accuracy, we plot the performance measure of relative error in Fig. 6c. The algorithm k2-zero (sdp) indeed has much higher accuracy for all instances. On the other hand, the algorithm k2-zero (ppa) also performs consistently and achieves relative errors below the threshold of  $\epsilon = 10^{-6}$  for all instances, which is significant given that the proximal point algorithm only uses first-order information.

Next, we will run the algorithm k2-zero (ppa) for larger instances. We set  $m = n$  and vary it from 50 to 500 with  $r = 5$  and  $s = \lceil 1.10 \cdot d_r \rceil$ , where  $d_r = r(m + n - r)$ . Note that for  $m = 500$ , the decision matrix  $X$  has the size of  $500 \times 500$  and there are approximately 5500 dense linear constraints in  $X$ , which requires a substantial amount of memory for the representation of instance data and the execution of the algorithm. For these instances, the tolerance is set to be  $\epsilon = 10^{-4}$  for  $\frac{\|X^{s+1} - X^s\|_F}{\max\{\|X^s\|_F, 1\}}$ , which is more consistent with respect to different instance sizes. The maximum number of iterations is kept at  $N_{\max} = 100$  as before. For the proximal point algorithm, we set the tolerance  $\epsilon_p = 10^{-6}$  and the maximum number of iterations  $N_{\max}^p = 100$ . Given these tolerances, the algorithm converges after approximately 14 iterations for all instances with 12 and 16 as the minimum and maximum, respectively. The relative errors are approximately  $10^{-4}$ , which are shown in Fig. 7a for different matrix sizes. The computation time per iteration increases significantly from 2.5 to  $10^4$  seconds when the matrix size increases from  $m = 50$  to  $m = 500$  as shown in Fig. 7b.

## 5 Conclusion

We have proposed non-convex models based on the dual Ky Fan 2- $k$ -norm for low-rank matrix recovery and developed a general DCA framework to solve the models. The computational results show that the proposed approach achieves high recoverability as compared to the truncated nuclear norm approach by Hu et al. [10] as well as the bilinear approach proposed by Jain et al. [13], especially when the size of the linear map is small. We also demonstrate that the proposed proximal point algorithm is promising for larger instances.

**Acknowledgements** The authors would like to thank the anonymous referee for his/her helpful comments and suggestions.

**Data Availability Declaration** The datasets generated during the current study are available from the corresponding author on reasonable request.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## A Appendix: Proof of Prop. 5

We start with some preliminary lemmas about the dual Ky Fan 2- $k$ -vector norm, which can be written as follows:

$$\begin{aligned} \|\mathbf{y}\|_{k,2}^* &= \max_{\mathbf{x}} \mathbf{y}^T \mathbf{x} \\ \text{s.t. } \|\mathbf{x}\|_{k,2} &\leq 1. \end{aligned} \quad (40)$$

Since  $\|\cdot\|_{k,2}$  is a symmetric gauge function (and so is its dual norm), we can focus on  $\mathbf{y}$  that satisfies  $y_1 \geq \dots \geq y_n \geq 0$ . We start with some properties of the optimal solution  $\mathbf{x}^*$  of (40).

**Lemma 2** *Consider the optimization problem in (40) with  $\mathbf{y}$  satisfying the above conditions. There exists an optimal solution  $\mathbf{x}^*$  that satisfies the following properties*

- (i)  $x_1^* \geq \dots \geq x_n^* \geq 0$ .
- (ii)  $x_i^* = x_k^*$  for all  $i \geq k + 1$ .

**Proof** (i) Consider an optimal solution  $\mathbf{x}$  and assume there exists  $j$  such that  $x_j < 0$ . Construct a new solution  $\mathbf{x}^*$  from  $\mathbf{x}$  by inverting the sign of the  $j$ -th element, which means  $x_j^* = -x_j > 0$ . We have,  $\|\cdot\|_{k,2}$  is a symmetric gauge function, thus  $\|\mathbf{x}^*\|_{k,2} = \|\mathbf{x}\|_{k,2} \leq 1$  or  $\mathbf{x}^*$  is a feasible solution. In addition,  $\mathbf{y}^T \mathbf{x}^* = \mathbf{y}^T \mathbf{x} - 2x_j y_j \geq \mathbf{y}^T \mathbf{x}$ . Thus  $\mathbf{x}^*$  is also an optimal solution, which means there exists a nonnegative optimal solution if  $\mathbf{y} \geq \mathbf{0}$ .

Now consider an optimal solution  $\mathbf{x}$  and assume there exists  $i < j$  such that  $x_i < x_j$ . Construct a new solution  $\mathbf{x}^*$  from  $\mathbf{x}$  by swapping two  $i$ -th and  $j$ -th elements, which means  $x_i^* = x_j$  and  $x_j^* = x_i$ . Using the properties of the symmetric gauge function, we again can show that  $\mathbf{x}^*$  is a feasible solution. We also have,  $\mathbf{y}^T \mathbf{x}^* = \mathbf{y}^T \mathbf{x} + (x_i - x_j)(y_j - y_i) \geq \mathbf{y}^T \mathbf{x}$

since  $y_i \geq y_j$ . Thus  $\mathbf{x}^*$  is an optimal solution and we have prove that there exists an optimal solution that satisfies  $x_1^* \geq \dots \geq x_n^* \geq 0$  if  $y_1 \geq \dots \geq y_n \geq 0$ .

- (ii) Consider an optimal solution  $\mathbf{x}^*$  that satisfies  $x_1^* \geq \dots \geq x_n^* \geq 0$ . We then have:  $x_i^* \leq x_k^*$  for all  $i \geq k + 1$ . We can construct the solution  $\bar{\mathbf{x}}$  such that  $\bar{x}_i = x_i^*$  for all  $i = 1, \dots, k$  and  $\bar{x}_i = x_k^*$  for all  $i \geq k + 1$ . We have:  $\|\bar{\mathbf{x}}\|_{k,2} = \|\mathbf{x}^*\|_{k,2} \leq 1$  and  $\mathbf{y}^T \bar{\mathbf{x}} \geq \mathbf{y}^T \mathbf{x}^*$ . Thus  $\bar{\mathbf{x}}$  is also an optimal solution of the problem shown in (40).

□

Applying Lemma 2, we can reformulate the problem (40) as follows:

$$\begin{aligned} \max \quad & \sum_{i=1}^{k-1} y_i x_i + \left( \sum_{i=k}^n y_i \right) x_k \\ \text{s.t.} \quad & \sum_{i=1}^k x_i^2 \leq 1, \\ & x_k \leq x_i, \quad i = 1, \dots, k-1. \end{aligned} \quad (41)$$

Without the second constraint set, the optimal solution  $\mathbf{x}^*$  of the relaxation problem of (41) can be obtained by using Cauchy-Schwartz inequality:

$$x_i^* = \frac{z_i}{\|\mathbf{z}\|}, \quad \forall i = 1, \dots, k,$$

where  $\mathbf{z} = \left( y_1, \dots, y_{k-1}, \sum_{i=k}^n y_i \right)$  (we can assume  $\mathbf{z} \neq \mathbf{0}$ ). Thus if  $y_{k-1} \geq \sum_{i=k}^n y_i$ ,  $\mathbf{x}^*$  is also

the optimal solution of (41). We will prove that if  $y_{k-1} < \sum_{i=k}^n y_i$ , then the optimal solution  $\mathbf{x}^*$  of (41) satisfies  $x_{k-1}^* = x_k^*$ . The generalization of this statement will be shown later but in order to prove it, we first need the following lemma.

**Lemma 3** Consider the following optimization problem

$$\begin{aligned} \max \quad & a_1 x_1 + k a_2 x_2 \\ \text{s.t.} \quad & x_1^2 + k x_2^2 \leq b, \\ & x_1 \geq x_2. \end{aligned} \quad (42)$$

where  $a_1, a_2 \geq 0$ ,  $a_1^2 + a_2^2 > 0$ ,  $b > 0$ , and  $k \in \mathbb{Z}_+$ .

- (i) If  $a_1 \geq a_2$  then the optimal solution can be found by solving the relaxation problem by removing the second constraint.  
(ii) If  $a_1 < a_2$  then the optimal solution satisfies  $x_1^* = x_2^*$ .

**Proof** (i) Removing the second constraint, we can solve the relaxation using Cauchy-Schwartz inequality and the optimal solution satisfies the condition  $x_1 \geq x_2$  since  $a_1 \geq a_2$ .

- (ii) We have:

$$\frac{1}{k+1} (a_1 + k a_2) (x_1 + k x_2) = (a_1 x_1 + k a_2 x_2) + \frac{k}{k+1} (a_1 - a_2) (x_2 - x_1) \geq a_1 x_1 + k a_2 x_2.$$

The equality happens when  $x_1 = x_2$  since  $a_1 < a_2$ . Solving the problem (42) with  $a_1 = a_2 = \frac{a_1 + k a_2}{k+1} > 0$ , we obtain the optimal solution  $\mathbf{x}^*$  that satisfies  $x_1^* = x_2^*$ . Thus

$$a_1 x_1^* + k a_2 x_2^* = \frac{1}{k+1} (a_1 + k a_2) (x_1^* + k x_2^*),$$

which implies that  $\mathbf{x}^*$  with  $x_1^* = x_2^*$  is also the optimal solution of (42) with the original parameters  $a_1$  and  $a_2$ .  $\square$

The structural property of the optimal solution of (41) can now be stated as follows.

**Lemma 4** *If  $y_{k-i} < \frac{1}{i} \left( \sum_{j=k-i+1}^n y_j \right)$  for an arbitrary  $i, i = 1, \dots, k-1$ , then there exists an optimal solution  $\mathbf{x}^*$  of (41) that satisfies the condition  $x_j^* = x_k^*$  for all  $j \geq k-i$ .*

**Proof** We prove the statement by induction. Let  $i = 1$ , consider an optimal solution  $\mathbf{x}^*$  of (41) and the following optimization problem:

$$\begin{aligned} \max \quad & \sum_{i=1}^{k-2} y_i x_i^* + y_{k-1} x_{k-1} + \left( \sum_{i=k}^n y_i \right) x_k \\ \text{s.t.} \quad & x_{k-1}^2 + x_k^2 \leq 1 - \sum_{i=1}^{k-2} (x_i^*)^2, \\ & x_k \leq x_{k-1}. \end{aligned}$$

Applying Lemma 3, we obtain the optimal solution  $\bar{x}_{k-1} = \bar{x}_k$ . We also have:  $2\bar{x}_k^2 = (x_{k-1}^*)^2 + (x_k^*)^2$  and  $x_{k-1}^* \geq x_k^*$ . Thus  $\bar{x}_k = \bar{x}_{k-1} \leq x_{k-1}^*$  (all values are nonnegative), which means  $\bar{x}_{k-1} \leq x_j^*$  for all  $j \leq k-2$ . Therefore,  $(x_i^*, \dots, x_{k-2}^*, \bar{x}_k, \bar{x}_k)$  is also an optimal solution.

Now assume the statement is true for  $i < k-1$ , we will prove that the statement is true for  $i+1$ . We have:  $y_{k-i-1} < \frac{1}{i+1} \left( \sum_{j=k-i}^n y_j \right)$  and  $y_{k-i-1} \geq y_{k-i}$ ; therefore  $y_{k-i} < \frac{1}{i} \left( \sum_{j=k-i+1}^n y_j \right)$ . Thus there exists an optimal solution  $\mathbf{x}^*$  that satisfies  $x_j^* = x_k^*$  for all  $j \geq k-i$ . We consider the restricted optimization problem of (41):

$$\begin{aligned} \max \quad & \sum_{j=1}^{k-1} y_j x_j + \left( \sum_{j=k}^n y_j \right) x_k \\ \text{s.t.} \quad & \sum_{j=1}^k x_j^2 \leq 1, \\ & x_k \leq x_j, \quad j = 1, \dots, k-i-1, \\ & x_j = x_k, \quad j \geq k-i, \end{aligned}$$

which can be used to solve the original problem. It is equivalent to the following problem:

$$\begin{aligned} \max \quad & \sum_{j=1}^{k-i-1} y_j x_j + (i+1) \left( \frac{1}{i+1} \sum_{j=k-i}^n y_j \right) x_k \\ \text{s.t.} \quad & \sum_{j=1}^{k-i-1} x_j^2 + (i+1)x_k^2 \leq 1, \\ & x_k \leq x_j, \quad j = 1, \dots, k-i-1, \\ & x_j = x_k, \quad j \geq k-i. \end{aligned}$$

Similar to the case  $i = 1$ , we can construct the problem with two decision variable  $x_{k-i-1}$  and  $x_k$  from an optimal solution of the problem above and prove that the solution is  $\bar{x}_{k-i-1} = \bar{x}_k$ . Thus the statement is true for  $i + 1$ , which means the statement is true for all  $i$  by induction.  $\square$

*Proof of Prop. 5.* Without loss of generality, let consider  $\mathbf{y}$  that satisfies the condition  $y_1 \geq \dots \geq y_n \geq 0$ . According to Lemma 4, with  $i^*$  defined in (29), there exists an optimal solution  $\mathbf{x}^*$  of (40) that satisfies  $x_i^* = x_{k-i^*}^*$  for all  $i \geq k - i^* + 1$  and all values can be found by solving the following optimization problem with the Cauchy-Schwartz inequality:

$$\begin{aligned} \max \quad & \sum_{j=1}^{k-i^*-1} y_j x_j + (i^* + 1) \left( \frac{1}{i^* + 1} \sum_{j=k-i^*}^n y_j \right) x_{k-i^*} \\ \text{s.t.} \quad & \sum_{j=1}^{k-i^*-1} x_j^2 + (i^* + 1) x_{k-i^*}^2 \leq 1. \end{aligned}$$

Note that the index  $i = 0$  always belongs to the considered set of indices. Thus we have:

$$\|\mathbf{y}\|_{k,2}^* = \sqrt{\sum_{j=1}^{k-i^*-1} y_j^2 + \frac{1}{i^* + 1} \left( \sum_{j=k-i^*}^n y_j \right)^2}, \quad (43)$$

which concludes the proof given that the above formulation can be easily extended for an arbitrary vector  $\mathbf{y}$ .  $\square$

## References

- Argyriou, A., Foygel, R., Srebro, N.: Sparse prediction with the  $k$ -support norm. In: NIPS, pp. 1466–1474. (2012)
- Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imag. Sci.* **1**, 183–202 (2009)
- Bhatia, R.: Matrix Analysis, Graduate Texts in Mathematics, vol. 169. Springer-Verlag, New York (1997)
- Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. *Found. Comput. Math.* **9**(6), 717–772 (2009)
- Candès, E.J., Tao, T.: Decoding by linear programming. *IEEE Trans. Inform. Theory* **51**(12), 4203–4215 (2005)
- Doan, X.V., Toh, K.C., Vavasis, S.: A proximal point algorithm for sequential feature extraction applications. *SIAM J. Sci. Comput.* **35**(1), A517–A540 (2013)
- Doan, X.V., Vavasis, S.: Finding the largest low-rank clusters with Ky Fan 2- $k$ -norm and  $\ell_1$ -norm. *SIAM J. Optim.* **26**(1), 274–312 (2016)
- Giraud, C.: Low rank multivariate regression. *Electron. J. Stat.* **5**, 775–799 (2011)
- Grant, M., Boyd, S.: CVX: Matlab Software for Disciplined Convex Programming, version 2.0 beta. <http://cvxr.com/cvx> (2013)
- Hu, Y., Zhang, D., Ye, J., Li, X., He, X.: Fast and accurate matrix completion via truncated nuclear norm regularization. *IEEE Trans. Patt. Anal. Mach. Intell.* **35**(9), 2117–2130 (2013). <https://doi.org/10.1109/TPAMI.2012.271>
- Hu, Z., Nie, F., Wang, R., Li, X.: Low rank regularization: a review. *Neural Netw.* (2020)
- Jacob, L., Bach, F., Vert, J.P.: Clustered multi-task learning: a convex formulation. *NIPS* **21**, 745–752 (2009)
- Jain, P., Netrapalli, P., Sanghavi, S.: Low-rank matrix completion using alternating minimization. In: Proceedings of the 45th Annual ACM Symposium on Theory of Computing, pp. 665–674. ACM (2013)
- Lee, K., Bresler, Y.: ADMiRA: atomic decomposition for minimum rank approximation. *IEEE Trans. Inform. Theory* **56**(9), 4402–4416 (2010)
- Liu, Y.J., Sun, D., Toh, K.C.: An implementable proximal point algorithmic framework for nuclear norm minimization. *Math. Program.* **133**(1–2), 399–436 (2012)



16. Ma, T.H., Lou, Y., Huang, T.Z.: Truncated  $\ell_1$ – $\ell_2$  models for sparse recovery and rank minimization. *SIAM J. Imag. Sci.* **10**(3), 1346–1380 (2017)
17. Mohan, K., Fazel, M.: Iterative reweighted algorithms for matrix rank minimization. *J. Mach. Learn. Res.* **13**(1), 3441–3473 (2012)
18. Nguyen, L.T., Kim, J., Shim, B.: Low-rank matrix completion: a contemporary survey. *IEEE Access* **7**, 94215–94237 (2019)
19. Pham-Dinh, T., Le-Thi, H.A.: Convex analysis approach to d.c. programming: theory, algorithms and applications. *Acta Math. Viet.* **22**(1), 289–355 (1997)
20. Pham-Dinh, T., Le-Thi, H.A.: A d.c. optimization algorithm for solving the trust-region subproblem. *SIAM J. Optim.* **8**(2), 476–505 (1998)
21. Recht, B., Fazel, M., Parrilo, P.: Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* **52**(3), 471–501 (2010)
22. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton, NJ (1970)
23. Toh, K.C., Todd, M.J., Tütüncü, R.H.: Sdpt3—a matlab software package for semidefinite programming, version 1.3. *Optim. Methods Softw.* **11**(1–4), 545–581 (1999)
24. Yin, P., Esser, E., Xin, J.: Ratio and difference of  $\ell_1$  and  $\ell_2$  norms and sparse representation with coherent dictionaries. *Commun. Inform. Syst.* **14**(2), 87–109 (2014)
25. Yin, P., Lou, Y., He, Q., Xin, J.: Minimization of  $\ell_1 - \ell_2$  for compressed sensing. *SIAM J. Sci. Comput.* **37**(1), A536–A563 (2015)
26. Ziętak, K.: Subdifferentials, faces, and dual matrices. *Linear Algeb. Appl.* **185**, 125–141 (1993)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.